DOCUMENT RESUME

ED 268 946 HE 019 314

AUTHOR Hunziker, Celeste M.
TITLE Evaluation of Remedial Programs at UC Davis.
PUB DATE Mar 86
NOTE 11p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Advising; Basic Skills; *College Instruction; Economically Disadvantaged; Educationally Disadvantaged; Educational Objectives; Educational Quality; Evaluation Criteria; Higher Education; High Risk Students; *Individualized Instruction; *Outcomes of Education; *Program Evaluation; *Remedial Programs; Research Design; *Standards; State Universities
IDENTIFIERS *University of California Davis

ABSTRACT

Efforts at the University of California, Davis, (UC Davis) to develop standard evaluation models for remedial programs are described, and three UC Davis evaluation studies are considered. A standard evaluation model entails a formal orientation, a singular values perspective, and a primary audience of program funders and oversight agencies. The research questions and the overall goals are defined prior to evaluation by these parties, not by campus evaluators/staff. Evaluations are based on predetermined program goals, and these goals may be translated into easily measurable outcome variables, (e.g., academic performance and retention), in order to develop comparable data across campuses. The evaluation design is established beforehand, and the methodology is carefully controlled and duplicated on each campus to ensure comparability of results. Successful evaluations involve four criteria: (1) the evaluation clearly identifies and specifies program goals, (2) the evaluation identifies and addresses the concerns of its primary audience, (3) the evaluation develops appropriate measures of program outcomes as they relate to program goals; and (4) the information provides a sufficient basis for judging program worth. The three UC Davis evaluation studies concentrated on Individualized Study Programs. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED268946

EVALUATION OF REMEDIAL PROGRAMS AT UC DAVIS

Celeste M. Hunziker

Student Affairs
Research and Information
University of California, Davis
March 1986

NIE 019 3/4

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Celeste M. Hunziker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EVALUATION OF REMEDIAL PROGRAMS AT UC DAVIS

Introduction

One of the fastest growing areas of the college curriculum consists of courses designed to help students overcome underpreparation for University work. Variously labeled "remedial," "developmental," or "compensatory," these courses concentrate on developing the basic reading, writing and mathematics skills necessary for students to have a reasonable chance of success in college-level studies.

Remediation at the University of California is not limited to introductory course work, but extends from admission and orientation up to graduation. It moves beyond the curriculum into the co-curriculum, including student services and activities throughout each campus. It involves summer bridge programs, individual and group tutoring in basic skills, workshops on skills development, and special academic advising and counseling. Because remediation is relative--to an institution, to a student, to a course of study--each UC campus has taken a unique approach in designing remedial activities consistent with its own internal structure and management philosophy and adapted to the distinctive character and needs of its own student body.

Most remedial programs developed within the last 15 years coincided with and were in partial response to the admittance of larger numbers of nontraditional students. Learning assistance centers evolved from traditional campus counseling centers to serve the needs of special admission students. Although these learning centers currently provide services to all students, they concentrate heavily on remedial support.

The Learning Skills Center at UC Davis provides a network of remedial support services designed to strengthen the basic skills of students who do not meet the University's regular admission requirements or who are considered to be educationally disadvantaged. Certain of these programs have been the subject of extensive evaluation efforts. The following report briefly discusses evaluation efforts at the University and the move toward developing standard evaluation models for remedial programs, and describes three recent UC Davis evaluation studies in detail.

Evaluation of Remedial Programs at UC

Certain types of evaluation, such as accreditation reviews, are routine at the University. Evaluation of remedial programs, however, is not quite so commonplace. A 1981-82 survey of remedial efforts in postsecondary education conducted by the California Postsecondary Education Commission (CPEC, 1983) indicated that, although all eight UC campuses report a positive effect of their remedial efforts, their evaluation efforts are haphazard at best. Survey responses suggested that evaluation ranged from tallies of program use to an occasional long-range study of grades and persistence.

Although all campuses report evaluating their remedial programs, the analyses performed rely primarily on student, staff and faculty evaluations done intermittantly. It appears that remedial services on most campuses are evaluated in the same way as traditional campus programs (i.e., user satisfaction). Few special studies to evaluate the success of remedial programs can be found.

If faculty and student evaluations are sufficient for the regular curriculum, what is the rationale for recommending that remedial programs be evaluated any more rigorously? Interest appears to stem from three factors: 1) the debatable position of remedial programs within the University, 2) the recent growth of these programs and, 3) their uncertain future, at least at the University of California. CPEC (1983) has called for the University to develop a standard evaluation model to be used to evaluate remedial courses and services on each campus and has requested that campuses use common criteria for evaluating programs and a common vocabulary for describing their activities and results.

Criteria for Developing a Standard Evaluation Model

In response to CPEC's recommendation, what criteria could the campus use? What assumptions would it have to make in order to develop a standard evaluation model for remedial programs?

A standard evaluation model developed according to CPEC's recommendation would have a formal orientation, a singular values perspective, and one primary audience--program funders and oversight agencies. The research questions and the overall goals would be defined prior to evaluation by these parties, not by evaluators or program staff on each campus. Although campuses would presumably be free to respond to the information needs of other interested stakeholders, the costs of carrying out such tasks could effectively preclude responding to other interests.

The basis for evaluations would be predetermined program goals. Because of the desire to develop comparable data across campuses, these goals would most likely be translated into easily measurable outcome variables, such as academic performance and retention. The evaluation design would be established beforehand and the methodology employed would have to be carefully controlled and duplicated on each campus to ensure comparability of results. Intermediate results might not be fed back into the program because that communication could result in changes in service on one campus, which would influence comparability of results. The oversight agency would set up a reporting schedule and each campus would respond on a set timetable using common vocabulary. On the basis of these findings, the oversight agency would make recommendations to the Legislature about whether and where remedial programs are working at the University, and whether or not these programs should continue. The emphasis would be on outcomes measured against program intents.

Summative evaluations conducted within the context of a standard evaluation model could be judged by the following criteria:

1) Does the evaluation clearly identify and specify program goals?

2) Does the evaluation identity and address the concerns of its primary audience?

3) Does the evaluation develop appropriate measures of program outcomes as they relate to the program goals?

4) Does the information provided constitute a sufficient basis for judging program worth?

Because evaluations based on a standard framework start with a uniform set of questions and possibly uniform evaluation formats for all types of programs, evaluation efforts would not meet the criteria identified above, but instead result solely in question-oriented studies. Although such studies are superior to true evaluation studies in efficiency of methodology or technical adequacy of information, their results are of limited use in improving programs and too narrow in scope to constitute a sound basis for judging worth. Concentration on predetermined goals and intents, particularly given likely resource constraints for conducting evaluations, might leave unanswered questions about side effects, unmet needs of specific populations, participant reactions, and goals and interests of other groups of stakeholders.

However, the foresight and planning necessary to formulate a standard evaluation model, and the resulting campus discussions of elements to be included in the model and the purposes of evaluation, could lead to more extensive use of evaluation. Evaluation efforts might become a part of the overall planning process as remedial programs are proposed and structured. Evaluation of intended means and ends would prevent campuses from developing unrealistic expectations of both their programs and their evaluations.

Evaluations of Remedial Programs at UC Davis

Three recent evaluations of remedial programs at UC Davis have been organized and carried out along the lines of the model described above. This section will discuss whether following a somewhat standard format allowed for meeting the four criteria for successful evaluations identified above.

These studies, conducted by Student Affairs Research and Information, evaluated services offered by the UC Davis Learning Skills Center (LSC). The services are part of the Individualized Study Program, which services educationally and/or economically disadvantaged students. This program is of particular interest to administrators and faculty alike because it enables students to reduce their work loads below the 12 units required by the University to meet minimum progress. This intrusion into the academic sphere, as seen against a background of increasing concern over the numbers

of students needing remedial services, resulted in evaluation efforts beyond tallies of program services and student and staff evaluations.

The Individualized Study Program has three components: 1) a reduced study load option, 2) basic skills workshops and, 3) a tracking and academic advising system. It was implemented as a two-year pilot program in 1981-82. An interim evaluation was conducted after the first year (Tom, 1983). Based on recommendations from this interim effort, two further evaluations were conducted (Hunziker 1984, 1985). The interim evaluation and one of the final evaluations concentrated on the first two components; the remaining evaluation was of the third component, called the Early Warning System (EWS). A separate evaluation of this component was conducted because the EWS serves considerably more students and because there was interest on the campus in developing similar tracking and academic advising systems in the colleges.

These three evaluations were decision-oriented studies, conducted to provide information necessary to determine whether or not the program should be continued. The evaluations were therefore directed at higher level administrators. The basic evaluation question, typical of summative evaluations of this nature, was: Did the program achieve the goals--improvement in academic performance and retention--set out in the program prospectus?

Although there were two attempts to study one set of program services, and one additional attempt to study a totally different type of service, all three studies employed the same basic quasi-experimental evaluation design. These evaluations were organized around problem statements on the effects of the interventions (independent variables) on specific outcome variables (dependent variables). The outcome variables, in this case GPAs and retention statistics, were derived from the program's goals.

The rationale for using a quasi-experimental design stemmed from the concern for establishing a relationship[1] between treatment and student achievement. In addition, it was thought that the use of comparison groups would provide an objective measure of whether or not the program was improving the academic performance and retention of participants and, therefore, meeting its goals.

Each of the three evaluation reports begins with a discussion of the underlying problem the program was developed to address. They relied on earlier campus research, which indicated that students in the targeted population were more likely to leave the University in academic difficulty than were their regularly admitted cohorts. This discussion was followed by a description of program services and a statement of the overall program objectives.

The initial evaluation of the reduced study load option and basic skills workshops took the program goals from the prospectus. Although the

---

[1]Because the program was directed toward high risk students, program staff rejected random assignment of students to treatment groups.

program remained virtually unchanged during the second year, the goals presented in the other two evaluations were translated into a series of short term goals described as intermediate steps in reaching the overall program goals contained in the prospectus. For example, the overall program objective of improvement in academic performance was accompanied by questions of academic achievement in specific target courses. The inconclusive results of the initial evaluation encouraged successive efforts to formulate short-term, less global outcome measures. Also, negative reactions to the lack of conclusive results from the first evaluation led to more "realistic preconditions" for further efforts, in particular a primary focus on positive outcomes. Although the goals were adjusted in the succeeding evaluations, the methodology for the central analyses remained unchanged.

Students were analyzed in comparison groups: students who participated in the program were in one group; students who were invited to participate, but who chose not to do so, were in the comparison group. Because students were not assigned randomly to these groups, an analysis of how comparable these groups were prior to the intervention was conducted using demographic and academic performance data. Differences between the groups were discussed and the extent and direction of bias resulting from these differences were estimated.

At this point, the initial evaluation went immediately to an analysis of year-end outcomes. Finding no differences in overall academic performance as measured by third quarter cumulative GPA based on program status, the evaluation concluded that there was no evidence to indicate that the program was having an impact on academic performance. Limitations in the study design were discussed but, in the evaluator's opinion, were not sufficient to explain the lack of conclusive results.

There was no discussion of program implementation or levels of student participation, although the program allowed for considerable variation in the latter. There was also no discussion of retention. This ommission was not discussed, although it should not be assumed that similarity in academic performance between the comparison groups implies the same for retention rates.

The first evaluation also included a participant survey conducted by telephone late in spring quarter. The survey had a 53% response rate; apparently many students could not be reached and many refused to participate. It does not appear that the enthusiastic responses of those who did participate in the survey, although presented in the report, were given much weight in the evaluator's conclusions about the program.

The evaluation recommended further study of the program, suggesting that any long term effects from the intervention be evaluated and that the target population for the program be reassessed.

The second evaluation of the same services followed a similar format to the same end. Again, limitations in the quasi-experimental design were discussed, but the report concluded that use of the reduced study load option together with the substitution of basic skills development activities

did not have an impact on overall academic performance. This time retention rates of students in the treatment and comparison groups were provided, but no differences were shown.

The second evaluation included a discussion of program implementation. Instead of most students participating in the program during fall quarter as originally intended, most participated later in the academic year, after having experienced considerable academic difficulty. Although the program allowed students to participate for 12 hours a week, in practice they chose to do so for only 2 to 5 hours per week. It was suggested that the timing and level of student participation patterns might explain lack of impact on global outcome measures, such as third quarter GPA.

In addition to the comparison group analysis and the discussion of student participation patterns, the evaluation included a section on short-term tracking of participants and of their performances in target courses. Although positive results from this type of analysis could not be attributed conclusively to program participation, it was suggested that there was some relationship between hours of verbal basic skills work and success in concurrent or immediately succeeding English courses. Put another way, if students were taking this kind of developmental activity but failing their English course work, the evidence would have been pretty damning for this intervention. This type of analysis reflects the fact that very few service programs directly attack or have any effect on their ultimate objectives; evidence that a program is affecting the intermediate process related to its objectives is therefore important.

The evaluation examined a possible negative side effect of allowing students to divert from the established curriculum. It concluded that, although participants did not take 12 units for one quarter, there were no differences in the number of minimum progress units completed at the end of three and six quarters based on program status. It appears that use of the reduced study load option does not impede students from resuming full course loads in succeeding quarters.

The evaluation also addressed the question of why the program appeared to have so little impact on outcome measures. An examination of participants and comparison students on a quarter-by-quarter basis suggested that program outcomes are dependent upon when students receive the intervention and the extent of academic difficulty they are in when they enter the program. Fall participants outperformed their comparison group; winter and spring quarter participants did not.

This evaluation recommended that, because there was no abuse of the reduced study load option and because program staff wish to serve each student individually, use of the reduced study load option should continue on a limited case-by-case basis at the discretion of the Program Director. It was also recommended that, because there were data to suggest that students needed a stronger intervention in mathematics, math workshops should be offered for more than 2 hours per week. The final recommendation was that the intervention was inadequate for students in considerable academic

difficulty; consequently, the program should be offered on only a limited basis during winter and spring quarters.

The third evaluation was of the Early Warning System. The format followed was similar to that of the second evaluation. Students were analyzed in comparison groups and end-of-year outcomes were compared. Again, there were no differences in overall academic performance, although this time there were differences in retention rates. Students who participated in the system were 15 to 20% more likely to return to UC Davis for a second year. Because students were not randomly assigned to groups, analyses were done to determine whether or not the EWS was contributing in some way to this result.

Extensive interview notes were kept, which permitted the follow-up of each piece of academic advice. Where possible, the academic outcome of each unit of advice was assessed (for example, a recommendation for tutoring in a specific course: Was it followed? Did the student pass the course?). This analysis indicated that students were likely to follow the advice they received and that the results were likely to be positive. Thus, the system appears to be a positive intermediate step toward improving the retention rates of this population. There were no recommendations attached to this evaluation.

Conclusion

All three evaluations were conducted within the constraints of the standard evaluation model discussed earlier. For the most part, they were successful in meeting three of the four evaluation criteria identified above and provide important information on the kind and level of benefits to expect from support programs of this nature. Limitations in study design did not preclude development of data to reduce, if not eliminate, decision makers' uncertainty about program outcomes. Goals of the program were clearly defined in these evaluations and the concerns of their primary audience identified and addressed.

The evaluations developed appropriate measures of intended program outcomes. The initial evaluation found that the program had no effect on these measures. After reaching the same conclusion, later evaluations developed separate measures and analyses of specific program activities. These analyses demonstrated that certain program services had a positive effect on the intermediate processes linking program services to intended goals, but that the intervention overall was insufficient to directly achieve those goals.

However, the information provided in these evaluations did not provide a sufficient basis for judging overall program worth. Effects from non-academic program services, such as students' increased commitment to the University or a reduction in their concerns, were not examined. Failure to address side effects was caused, in part, by the attachment to a particular methodology. Each evaluation referred to the non-academic program services,

but eschewed their evaluation in an attempt to get more easily measurable bottom-line results.

The evaluations concentrated on measuring outcomes against program goals and were of limited use in improving the program. For example, although the second and third evaluations contained important formative information on program services, this information was not fed back into the program as it was developed. Without this kind of detailed and timely information on program implementation and specific service outcomes, managers will find it difficult to modify their programs or to adopt programs successfully from other campuses. This type of formative information is important, particularly because findings of "no program effects" in the past have not been used by decision makers to eliminate programs. The need for summative go/no go information has never been clearly demonstrated.

Based on evaluation efforts to date, it appears that other evaluation strategies may be of more use to all parties. These efforts should be organized around such questions as: Given the nature of the client group, what effects can be realistically expected from the program? To what extent is the program theory a valid one? Will the service logically produce the effects it claims? Despite the change in focus, process evaluations of this type could still provide information to judge whether specific components of a program were working.

In a time of increasingly scarce resources for program staff and evaluators alike, it is necessary to question whether concentrating on the same evaluation approach will yield anything more useful in the future than it has in the past. It is not clear that it would. Ultimately we may have to admit that some students need far more support than the University is willing or able to provide. In the meantime, more of the same kinds of evaluation may not be a helpful or productive way of spending our time and money.

REFERENCE LIST

California Postsecondary Education Commission. Promises to Keep. Sacramento, 1983.

Hunziker, Celeste M. Evaluation of the Individualized Study Program: Early Warning System. Student Affairs Research and Information, University of California, Davis, 1984.

Hunziker, Celeste M. Evaluation of the Individualized Study Program: Reduced Study Load Option and Basic Skills Workshops. Student Affairs Research and Information, University of California, Davis, 1985.

Tom, Alice K. Interim Evaluation Report: Individualized Study Program. Student Affairs Research and Information, University of California, Davis, 1983.

paper1.cmh